Sun-Pharma Case Study: Assignment

Question 1:

Quality assurance checks on the previous batches of medications found that it is four times more likely that a medicine is able to produce a satisfactory result than not.

Given a small sample of 10 medicines, you are required to find the theoretical probability that, at most, 3 medicines are unable to do a satisfactory job:

1) Propose the type of probability distribution that would accurately portray the above-mentioned scenario, and list out the three conditions that this distribution follows.

Ans:

Based on the scenario provided, the type of probability distribution that would accurately portray this situation is the ***binomial distribution***. The binomial distribution has the following three conditions:

1. The trials are independent: Each medicine's ability to produce a satisfactory result is independent of the others.

2. The probability of success is constant: The probability of producing a satisfactory result is 4 times more likely than not, which means the probability of success is 4/5 or 0.8.

3. There are a fixed number of trials: In this case, we are testing 10 medicines.

2) Calculate the required probability.

Ans.

Let, X – the no. of drugs that produce unsatisfactory result after testing total 10 drugs.

Now, X will be =10.

n=>10.

P – the probability for unsatisfactory result.

So, the probability will be given by P (unsatisfactory result) +4P(4 times more likely that a drug is able to produce a satisfactory result).

Therefore, P+4P=1

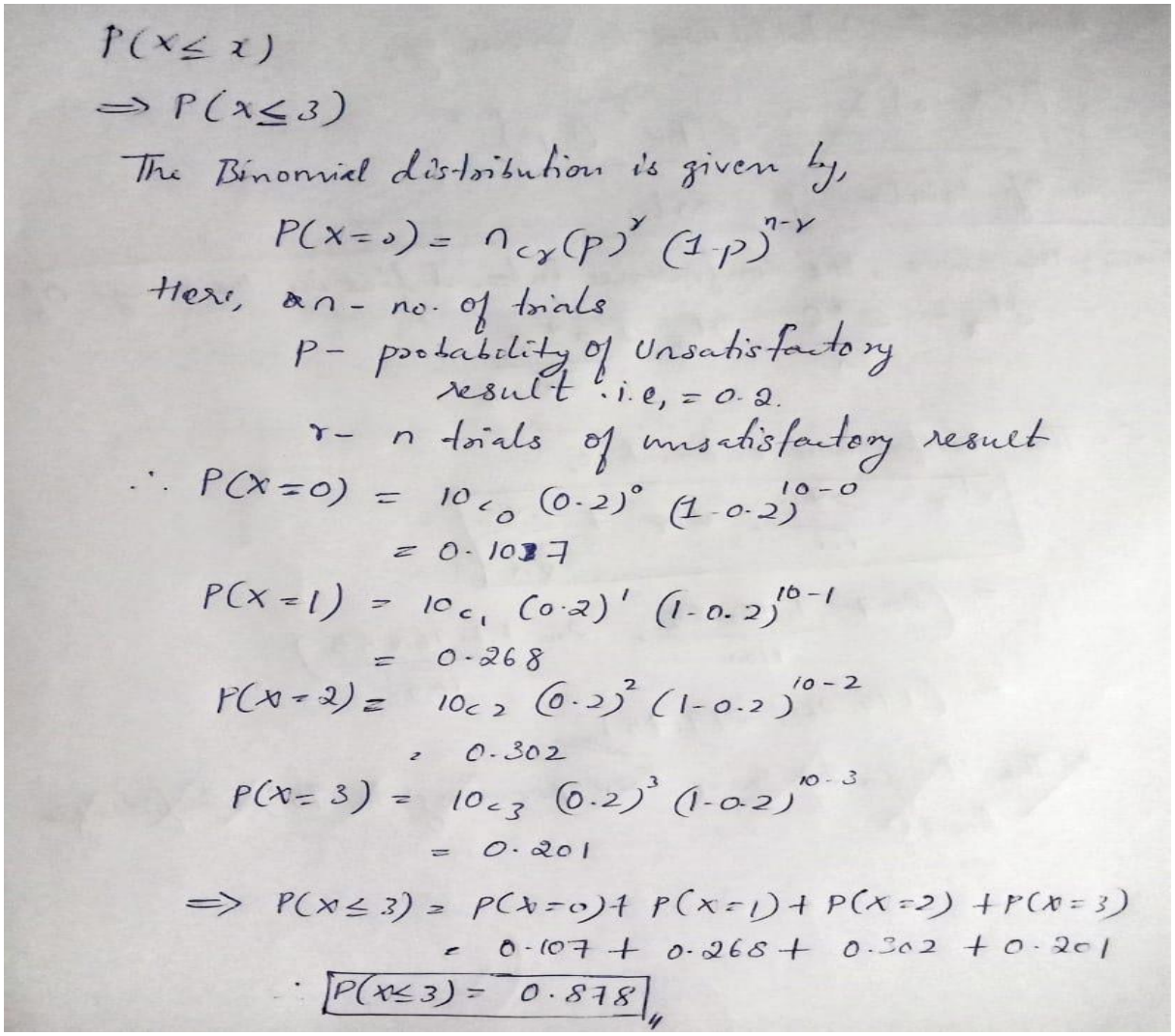
5P=1,

P=0.2

Probability that drug produces unsatisfactory result =0.2

For Satisfactory result=1-0.2=0.8.

Now, for theoretical probability for at most 3 times, that the drug not able to do a satisfactory job can be defined as cumulative probability of X. Given as,



Hence, (X ≤3) =0.878= 87.8%.

Therefore, the probability that at most 3 drugs are unable to do satisfactory job is *87.8%.*

Question-2:

For the effectiveness test, a sample of 100 medicines was taken. The mean time of effect was 207 seconds, with the standard deviation coming to 65 seconds. Using this information, you are required to estimate the interval in which the population mean might lie – with a 95% confidence level:

1) Discuss the main methodology using which you will approach this problem. State all the properties of the required method. Limit your answer to 150 words.

Ans:

The main methodology using which we will approach this problem is ***Central Limit Theorem.***

The Central Limit Theorem consists of three properties such as,

- Sampling Distribution mean=Population mean.
- Sampling distribution's standard deviation **(Standard error)** = $\boldsymbol{\sigma}$ / √$\boldsymbol{n}$, where σ is the population's standard deviation and n is the sample size.
- For n > 30, the sampling distribution becomes a normal distribution.

Let's Consider, that the sample mean would be normally distributed with mean (μ) ,which allows us to calculate the standard deviation of the sampling distribution, (S) is the sample's standard deviation and (σ) is population's standard deviation. Furthermore, in the particular scenario, we only know the sample's standard deviation.

2) Find the required interval.

Ans.

n (Sample Size) = 100

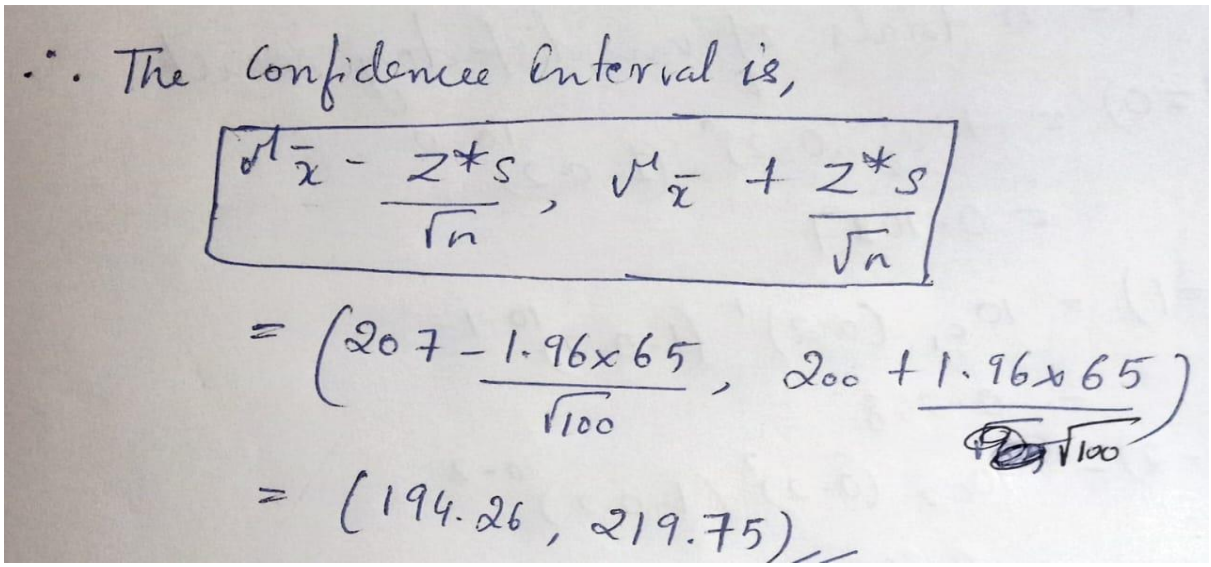
$\mu\bar{x}$ (Sample mean) = 207

S (Sample standard deviation) = 65.

Here $Z*$ is the Z-score associated with a confidence level of 95%.

Z* Values for Commonly Used Confidence Levels are:

Confidence level	Z*
90%	± 0.65
95%	± 1.96
99%	± 2.58



Therefore, the population mean (μ) lies between *194.26 seconds and 219.75 seconds.*

Question-3:

1) The painkiller needs to have a time of effect of at most 200 seconds to be considered as having done a satisfactory job. Given the same sample data (size, mean and standard deviation) as that in the previous question, test the claim that the newer batch produces a satisfactory result and passes the quality assurance test. Utilise two hypothesis testing methods to take a decision. Take the significance level at 5%. Clearly specify the hypotheses, the calculated test statistics and the final decision that should be made for each method.

Ans.

Null(H0): H0: $\mu \leq 200$ seconds, considered as having done a satisfactory job.

H1: $\mu > 200$ seconds, considered as not having done a satisfactory job.

The type of test is ***One-tailed test (Upper tailed test),*** since the its one directional hypothesis and rejection side will be on the right side.

Given, Sample size n = 100

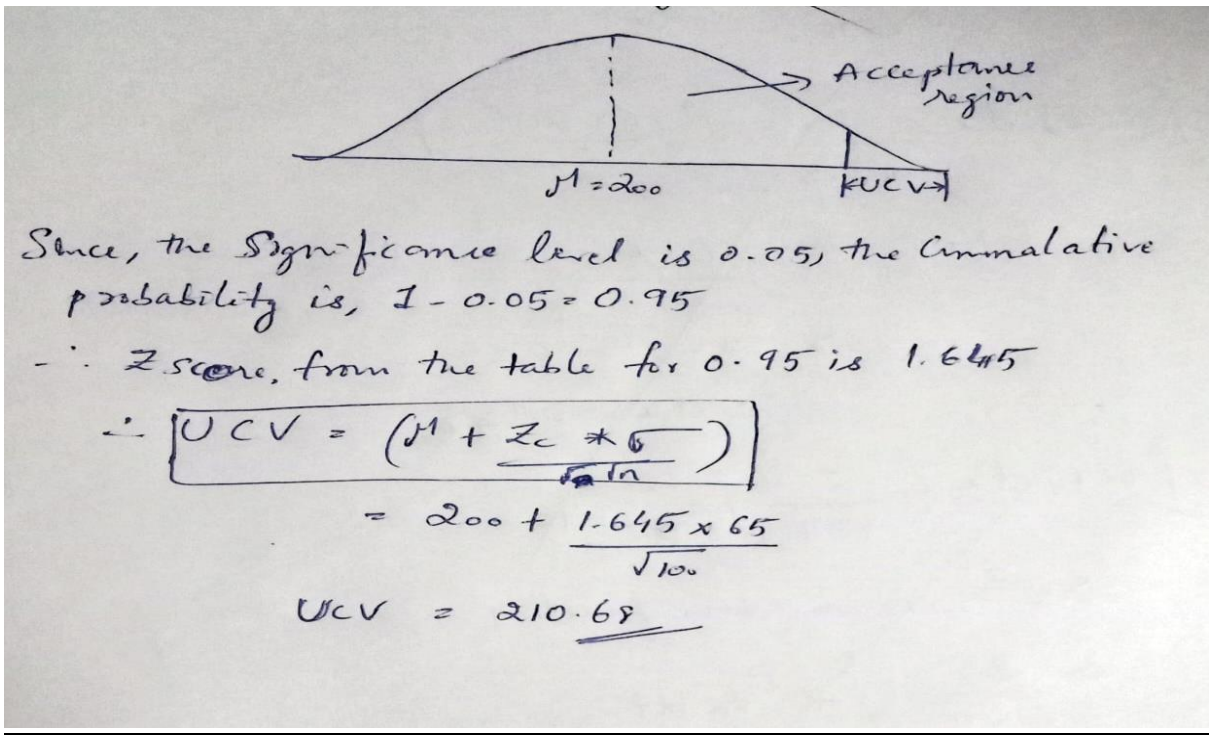
Assumed Sample mean (μ) = 200

Sample mean ($\mu\bar{x}$)= 207

Sample standard deviation ($\sigma\bar{x}$) = 65.

Significance level α = 5% = 0.05

Critical Value Test:



Since, the Significance level is 0.05, the Cummalative probability is, 1 - 0.05 = 0.95

∴ Z score, from the table for 0.95 is 1.645

$$\therefore \boxed{UCV = \left(\mu + Z_c * \frac{\sigma}{\sqrt{n}}\right)}$$

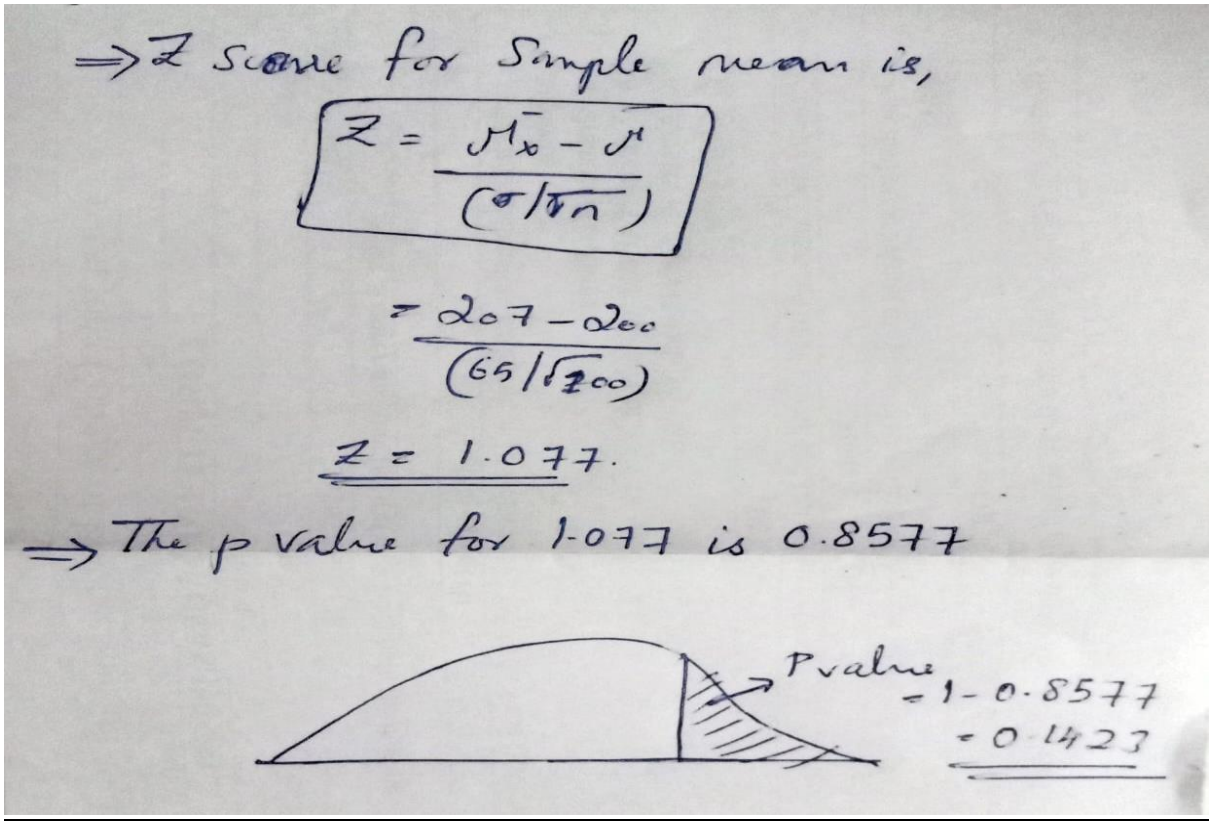
$$= 200 + \frac{1.645 \times 65}{\sqrt{100}}$$

$$UCV = 210.68$$

Since the sample mean ($\mu\bar{x}$) =207 seconds is less than the Upper Critical Value of 210.68 seconds, ***it lies in the acceptance region***.

Therefore, we *fail to reject the null hypothesis.* It specifies that for a medicine to be considered to have performed satisfactorily, the time of effect must be at most 200 seconds.

p-value Test:



Since the p-value (0.1423 > 0.05), we *fail to reject the null hypothesis* which states that the drug needs a time of effect at most 200 seconds to do a satisfactory job.

2) You know that two types of errors can occur during hypothesis testing – Type I and Type II errors – whose probabilities are denoted by α and β, respectively. For the current sample conditions (sample size, mean and standard deviation), the value of α and β come out to be 0.05 and 0.45, respectively.

Now, a different sampling procedure (different sample size, mean and standard deviation) is proposed so that when the same hypothesis test is conducted, the values of α and β are controlled at 0.15 each. Under what conditions would either method be more preferred than the other? Give an example of a situation where conducting the hypothesis test with α and β as 0.05 and 0.45, respectively, would be preferred over conducting the same hypothesis test with α and β at 0.15 each. Similarly, give an example for the reverse scenario, where conducting the same hypothesis test with α and β at 0.15 each would be preferred over having them at 0.05 and 0.45, respectively. For each example, give suitable reasons for your

particular choice using the given values of α and β only. (Assume that no other information is available. Additionally, the hypothesis test that you are conducting is the same as mentioned in the previous question; you need to test whether the newer batch produces satisfactory results.)

Ans.

Type-1 (α) error refers to **reject the NULL hypothesis** even it is true.

Type-2 (β) error refers to the scenario where we **failed to reject the NULL hypothesis** when it was false.

The possible scenarios are:

- The value of α and β come out to be 0.05 and 0.45 respectively.
- The value of α and β are at 0.15 each.

When there are negative consequences associated with accepting an alternate hypothesis, we work to keep the error to a very low value. We need to reject the null hypothesis conservatively because, based on the circumstance in question, if there is any adverse effect connected to the overdose of the painkiller drug, we need to minimise the error to a minimal. Therefore, we would like to maintain the value of and, which equals 0.05 and 0.45, respectively, if the painkiller drug has substantial adverse effects. In contrast, if the painkiller drug had no significant adverse effects, we would be more accepting of the inaccuracy. This would imply that we can actively challenge the current situation and develop the alternative hypothesis.

Question 4:

Once one batch has passed all the quality tests and is ready to be launched in the market, the marketing team needs to plan an effective online ad campaign to attract new subscribers. Two taglines were proposed for the campaign, and the team is currently divided on which option to use.

Explain why and how A/B testing can be used to decide which option is more effective. Give a stepwise procedure for the test that needs to be conducted.

Ans.

A/B testing is a statistical method used to compare two different versions of a marketing campaign to determine which one is more effective. In this case, A/B testing can be used to decide which of the two taglines proposed by the marketing team is more effective in attracting new subscribers for the medication.

A/B testing can be particularly useful in marketing as it enables companies to optimize their campaigns for better performance and improve their return on investment (ROI). By testing different versions of a campaign or messaging, companies can determine which version resonates better with their target audience and adjust their strategy accordingly. This can

result in higher engagement rates, increased conversions, and ultimately, increased revenue. A/B testing allows for maximum output with minimal modifications, eventually leading to increased revenue and substantial growth.

Stepwise procedure for conducting an A/B test:

- **Define the objective:** Clearly define the objective of the A/B test. For example, the objective could be to increase the number of subscribers for the medication.
- **Define the variables:** Identify the variables that will be tested. In this case, the variables would be the two different taglines.
- **Randomly divide the audience:** Randomly divide the target audience into two groups, one for each tagline. It is important to ensure that the two groups are similar in all other characteristics except for the tagline they see.
- **Show the taglines:** Show each group the respective tagline. This can be done by creating two different landing pages or ads with the different taglines and randomly assigning the audience to the two groups.
- **Measure the response:** Measure the response rate for each group. This can be done by tracking the number of subscribers or clicks generated by each tagline.
- **Analyze the data:** Analyze the data using statistical methods such as hypothesis testing to determine whether there is a statistically significant difference between the two groups. The level of significance and the statistical test used will depend on the sample size, data distribution, and other factors.
- **Draw conclusions:** Based on the analysis, draw conclusions about which tagline is more effective. If there is a statistically significant difference between the two groups, choose the tagline with the higher response rate.
- **Implement the winning tagline:** Once the winning tagline is identified, implement it in the marketing campaign.

By using this method, the marketing team can make data-driven decisions and optimize the ad campaign for better performance, which can ultimately lead to increased subscribership and revenue.